Journal of Computer Science and Electrical Engineering
Print ISSN: 2663-1938
Online ISSN: 2663-1946
DOI: https://doi.org/10.61784/jcsee3094

DEEP LEARNING DRIVEN ANALYSIS OF THE FOOD-RELATED VIDEO COMMUNICATION EFFECT: INVESTIGATING VISUAL FEATURE IMPACTS

YaWei He*, ZiJing Pan, LinYi Bao, YanXi Zhu, ChiYue Zhang, ZhouChu Zhang
School of Media and Communication, Shanghai Jiao Tong University, Shanghai 200240, China.
Corresponding Author: YaWei He, Email: heyawei1992@sjtu.edu.cn

Abstract: To extend the scope of computational communication in an era dominated by visual media, this study explores effective methods for analyzing video content at scale, moving beyond a traditional focus on textual data. Using a large dataset of food-related videos from the popular platform Bilibili, we developed and trained a custom deep learning model to automatically identify, extract, and quantify the presence of core visual elements within each video frame. A systematic correlation analysis was conducted to examine the relationship between these extracted visual features—specifically the face appearance rate, food appearance rate, and overall image brightness—and composite measures of the videos' communication effects. Our statistical analysis reveals that both a higher face appearance rate and greater image brightness are significant positive predictors of communication effectiveness. In contrast, and counter to common assumptions, a higher frequency of shots featuring only food was found to negatively impact the video's overall performance. These findings suggest that effective video communication relies on emotional connection rather than mere content display; facial presence significantly drives deep engagement, likely through social relationships, while the visibility of food itself negatively impact audience response, highlighting a preference for cultural context and human narratives. Furthermore, metadata such as expressive titles and channel fan count, along with release time and duration, also critically shape dissemination success. This research not only offers valuable empirical guidance for content creators but also demonstrates a replicable and cost-effective computational paradigm for large-scale video content analysis.
Keywords: Artificial intelligence; Deep learning; Computational communication; Communication effect; Visual features

1 INTRODUCTION

The natural instinct of humans to pursue visual expression has led to a shift in the main source of information acquisition from text and images to videos. The "China Online Audio - Visual Development Research Report (2024)" shows that as of December 2023, the scale of online audio - visual users had reached 1.074 billion, accounting for 98.3% of the total number of Internet users. Online videos have become an important form of entertainment in people's daily lives. The rapid development of online videos has provided massive video data for research based on them. At the same time, thanks to mature communication research methods, the research on the communication effects of various online videos in the field of communication has formed a certain scale and system [1-2]. With the rapid development of artificial intelligence technology, deep learning algorithms and computing power have enabled the continuous development of computational communication [3-8], which has significant interdisciplinary integration characteristics.

In terms of the research content of computational communication, it focuses on the content of "communication". Zhijin Zhong and other scholars have compared Chinese and English computational communication studies from three dimensions: research topics, methods, and theories, they pointed out that in terms of research topics, political communication is a common concern of scholars in China and the West [9]. In addition, the research content of computational communication mainly involves fields such as online public opinion [10], news production [11], computational propaganda [12], health communication and media use [13-14], showing strong disciplinary openness. From the perspective of computational communication research methods, Zhijin Zhong and others found that semantic analysis is the most commonly used analysis method in computational communication research in the past year, along with network analysis and time series analysis, and a new trend of combining computational methods with other quantitative or qualitative research methods is emerging [9]. The continuous expansion of computational communication research content on communication topics and the innovation of using new computational methods are driving the continued development of computational communication [15].

However, most studies are still focus on the text part of various communication topics [16]. Only a small number of video - based studies are often limited to the language and text in the videos, and relatively few studies focus on the main video content and video style feature variables [17]. This is mainly because there are certain technical difficulties in extracting such video feature variables. With the help of computer vision technology, the extraction of aesthetic features of images, portrait recognition technology, and object feature recognition functions can be down [18]. However, in terms of the technical tools used, the vast majority of studies use paid APIs from companies such as Google and Microsoft [19]. For example, Wu Ye et al. used the Face++ API to analyze the visual content such as the gender, expression, gestures, and head posture of the bloggers in the "News Anchors on the Beat" column on Douyin, verifying the personalized communication effect of mainstream media short videos [20]. However, when conducting large - scale research using APIs, certain costs will incurred, and the research may limited by the model.

This study selects food - related videos on Bilibili as the research objects. These videos exist in large numbers in self - media. The video types themselves involve the intersection of multiple communication fields such as urban communication, cross - cultural communication, and network communication, and there are differences in video communication effects [21]. In terms of research methods: First, this study uses the Python3.8.5 virtual environment as the basic environment for code operation. Through programming crawlers, it obtains food - related videos on Bilibili, ensuring that the sample is representative and complete. Second, this study innovatively uses the existing deep - learning algorithm YOLOv5 by means of independent annotation and model training, it realizes the automated large - scale identification of the main content in videos, and for food - related videos, we proposed two main video content feature indicators, namely the face appearance rate and the food appearance rate. Finally, this study combines open - source databases such as OpenCV to extract video feature indicators, video metadata indicators, and communication effect indicators under the visual frame theory, and analyzed the correlation between the extracted video indicators and the communication effect through regression algorithms. The research conclusions we found provide a new direction for self - media video creators to optimize video content in actual creation.

2 SELECTION OF RESEARCH VARIABLES AND FORMULATION OF QUESTIONS

This project based on the mature visual frame theory proposed by Rodriguez&Dimitrova [22]. Taking food - related videos on Bilibili as the research objects, we extracts four video features that reflect the main content of the videos: face appearance rate, food appearance rate, brightness, and image entropy.

(1) Face Appearance Rate: Facial recognition and face detection are among the most widely used applications in computer vision technology. As the human face is a major subject that appears in videos, we are curious about the impact of its appearance frequency on the video communication effect. Therefore, Hypothesis H1 is proposed.

(2) Food Appearance Rate: In food - related videos, apart from people, food is the main content of the video. An anchor will display and introduce the quality, production process, and sensory experience of dishes based on preselected restaurants, eateries, or street food stalls. The video often uses wide-angle or tight shots to show either the abundance of food varieties or the appealing beauty of the dishes. It is the wish of food - related video content producers to make the audience feel the "shared eating" experience through the videos. Therefore, we set the food appearance rate as a video feature indicator reflecting the main content and study its impact on the video communication effect. Hypothesis H2 is proposed.

(3) Brightness: Brightness considered as fundamental to human visual perception [23]. We wonder whether the difference in image brightness plays a role in attracting the audience's attention, affecting visual perception memory, and influencing the communication effect. Thus, Hypothesis H3 is proposed.

(4) Image Entropy: The concept of entropy comes from information theory, which refers to the inherent uncertainty in the possible results of a random variable [24]. In the field of images, entropy conceptualized as the heterogeneity of pixels in an image, reflecting the amount of information contained in the image. Higher entropy implies more information and finer details per unit area. We want to know the impact of image entropy on the video communication effect and propose hypothesis H4.

Based on the above, we propose the following four hypotheses:

- H1: In food - related videos, the face appearance rate has a significant positive impact on the video communication effect.
- H2: In food - related videos, the food appearance rate has a significant positive impact on the video communication effect.
- H3: In food - related videos, the brightness has a significant positive impact on the video communication effect.
- H4: In food - related videos, the video entropy value has a significant positive impact on the video communication effect.

In addition to the four video feature indicators involved in the above visual framework, this study also collects video metadata variables: video duration, release time, whether there are question marks and exclamation marks in the title, and relevant data of the video bloggers - the number of blogger's fans. The number of fans used as an important indicator to measure the influence of video bloggers. It is worth exploring what differences exist in the creation of videos on the same subject by video bloggers with different numbers of fans. Therefore, the following research question Q1 and four hypothesis based on feature indicators H5-H8 are proposed:

- Q1: Do the visual frame features of food - related videos by Bilibili video bloggers with different numbers of fans differ? If so, what are the differences?
- H5: In food - related videos, the title punctuation has a significant positive impact on the video communication effect.
- H6: In food - related videos, the release time has a significant positive impact on the video communication effect.
- H7: In food - related videos, the video duration has a significant positive impact on the video communication effect.
- H8: In food - related videos, the video fan count has a significant positive impact on the video communication effect.

3 RESEARCH METHODS

3.1 Data Source

This study selects videos in the "Food Detective" section under the "Food Area" of Bilibili as the research objects. Through Python crawlers, we captured 1000 food - related videos. The video release time ranges from February 2019 to August 2023, and the playback volume ranges from 2.38 million to 16.25 million. After screening the integrity of data crawling and downloading, a total of N=889 videos were finally retained as the sample dataset for this study.

3.2 Video Feature Extraction

3.2.1 Face appearance rate and food appearance rate

Since the essence of a video is a continuous sequence of picture frames, the appearance rates of faces and food in video images are essentially the ratios of the total counts of face and food elements in single - frame pictures to the total length of the video. First, this study uses the OpenCV open - source package to read all consecutive frames of the video set, and automatically obtains 72,208 pictures through a program. Then, this study uses the deep - learning Yolov5 algorithm to train a model and uses the trained model to identify the faces and the food in the above - obtained picture set automatically. To obtain the training model, we completed the following three steps: (1) we divided the dataset was into a training set and a test set, and the training set data was manually annotated for faces and food; (2) The Yolov5 algorithm was used to train the annotated training set data to obtain a model. The training results indicate that both the precision (P) and recall (R) exceed 0.85 (see Table 1), and the loss function curve consistently exhibits a fluctuating decline (see Figure 1), demonstrating effective training outcomes for our model. (3) Feed the test set into the trained model to obtain detection results for faces and food. Some face and food detection results show in Figure 2. Finally, were calculated the appearance rates of the faces and the food in each video cumulatively, and the face appearance rate and food appearance rate values were obtained by taking the natural logarithm of the calculation results.

Table 1 Performance Evaluation of Yolov5 Training Model

Class	Images	Labels	P	R	mAP@0.5
All	100	151	0.858	0.877	0.921
Face	100	64	0.846	0.984	0.967
Food	100	87	0.87	0.769	0. 875

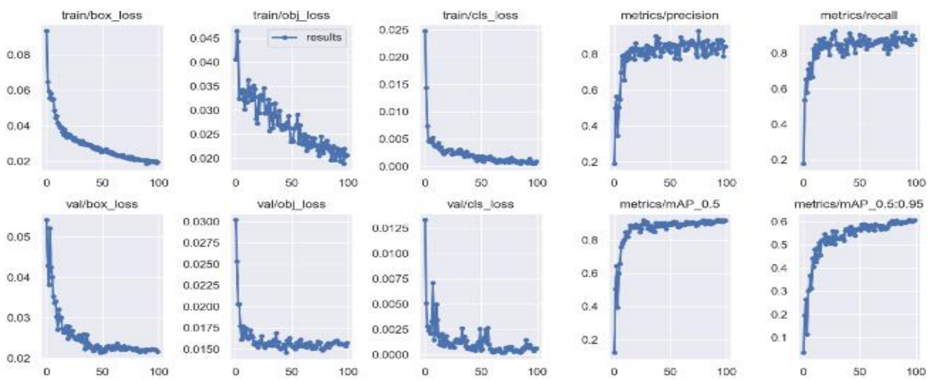

Figure 1 Performance Evaluation of Yolov5 Model

Figure 2 Automatic Face and Food Detection by Deep Learning Yolov5 Algorithm

3.2.2 Information entropy

This study measures image entropy through the Shannon entropy formula and its calculation formula is as Equation 1. This study first uses the OpenCV library of the Python program to convert all sampled frames into grayscale images to reduce the image information dimension and improve the operation processing speed.

In the formula, P(x) represents the proportion of pixels with gray value i (i=1,,,,n) in the image. The entropy value of a frame with all - black pixels will be 0, while a highly textured frame will have a higher entropy value (with a maximum of no more than 8). This study automatically obtains the average value of all sampled frames in each video through code, and finally obtains the image entropy values of all videos in this study through data standardization processing.

$$H(\mathrm{X}) =- \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{P}(\mathrm{X_i}) \log_{\mathrm{b}} \mathrm{P}(\mathrm{X_i}) \quad (1)$$

3.2.3 Image Brightness

HSV is a method of representing points in the RGB color space in an inverted cone, HSV stands for Hue, Saturation, and Value, also known as HSB (B stands for Brightness). This study uses the cv2 library of Python to collect the HSV data of the video and extracts the value component as the image brightness value. The brightness values of all sampled frames of each video are calculate arithmetically by code after collect the brightness of the video image.

3.2.4 Video Metadata Extraction

This study also extracts video metadata such as video duration, release time, whether there are question marks and exclamation marks in the title, whether there are numbers in the title, and the number of fans of the video account. The video duration is in seconds. The release time designed as a binary variable of "newly released" and "previously released" with January 2023 as the boundary. Video metadata such as whether there are question marks and exclamation marks in the title and whether there are numbers in the title are all set as binary variables of 0 and 1, with 1 marked if present and 0 marked if absent. According to the number of fans of the account, video accounts defined into four levels: less than 100,000 fans, 100,000 - 1,000,000 fans, 1,000,000 - 5,000,000 fans, and more than 5,000,000 fans manually.

3.3 Communication Effect Measurement

The measurement for the communication effect of video (noted by letter C) in this study included the following four dimensions (all measurement units are in ten thousand). First, we use the video play count to measure "communication breadth" (noted by letter B). Second, we use the number of video coins to measure "communication approval" (noted by letter A). Third, we use the number of video comments to measure "communication participation" (noted by letter P), and then we use the number of video forwards to measure "communication sociability" (noted by letter S). We also used OpenCV to extract the communication effect indicators of each videos automatically. We assign weights of 0.5, 0.3, and 0.2 to communication breadth, approval, and sociability, as well as participation respectively. After taking the natural logarithm of the results, we can get the communication effect value. The calculation formula is as Equation 2:

$$C = In(0.5B + 0.3(A + S) + 0.2P) \quad (2)$$

4 RESULTS

4.1 Analysis of the Difference between the Number of Fans of Food-Related Video Bloggers and the Visual Frame Features

To answer Q1, through the chi-square test, we found that in the food-related videos posted by video bloggers with different numbers of fans, the difference in the face appearance rate at different fan levels reached a significant level ($p<0.01$). In terms of different levels, the face appearance rate of videos by bloggers with more than 5 million fans is 23%, that of bloggers with 1-5 million fans is 15%, that of bloggers with 0.1-1 million fans is 12%, and that of bloggers with less than 0.1 million fans is 6%. It shows that the higher the number of fans, the higher the face appearance rate.

In the food-related videos posted by video bloggers with different numbers of fans, the difference in the food appearance rate at different fan levels also reached a significant level ($p<0.01$). In terms of different levels, the food appearance rate of videos by bloggers with more than 5 million fans is 9.18%, that of bloggers with 1-5 million fans is 12.27%, that of bloggers with 0.1-1 million fans is 9.90%, and that of bloggers with less than 0.1 million fans is 8.36%. The food appearance rate of bloggers with 1-5 million fans is higher than other groups.

For image brightness and image entropy values, there is no significant difference at different levels of the number of fans ($p>0.05$).

4.2 Regression Analysis of Video Features and Video Communication Effects of Food-Related Videos

To verify H1-H4, through regression analysis we get the results (see Table 2), we found that the face appearance rate has a significant positive impact on the communication recognition degree ($\beta=0.170$, $p<0.01$), communication participation degree ($\beta=0.18$, $p<0.01$) and communication social degree ($\beta=0.111$, $p<0.01$), so hypothesis H1 holds. The food appearance rate has a significant negative impact on the communication breadth ($\beta=-0.135$, $p<0.01$), communication recognition degree (β=−0.169, p<0.01), communication participation degree (β=−0.116, p<0.01), communication social degree (β=−0.021, p<0.01) and communication effect (β=−0.133, p<0.01), so hypothesis H2 does not hold. There is a significant positive impact between image brightness and the communication effect (β=0.072, p<0.05), so hypothesis H3 holds; there is no significant relationship between image entropy and the communication effect, so hypothesis H4 does not hold.

Regarding the video metadata variables and to verify H5-H8, through regression analysis we get the results (see Table 2), the analysis found that videos including a question mark or exclamation mark in the title achieved a positive effect on communication effect (β=0.181, p<0.01), so H5 holds. The release-time has a negative impact on the communication sociability (β=−0.029 p<0.01), so H6 does not hold. The video duration has a positive impact on the communication approval (β=0.26, p<0.01) and communication participation (β=0.303, p<0.01), so H7 hold. The video fan count has a positive impact on the communication approval (β=0.18, p<0.01) and communication participation (β=0.25, p<0.01), so H8 hold.

Table 2 Regression Analysis of Video Features and Video Communication Effects of Food - Related Videos

Variables	Communication Breadth	Communication Approval	Communicatio n Participation	Communication Sociability	Communication Effect
Video Features					
Face Appearance Rate	0.028	0.170***	0.180***	0.111***	0.018
Food Appearance Rate	-0.135***	-0.169***	-0.116***	-0.021***	-0.133***
Image Brightness	0.096**	-0.042	0.004	0.128***	0.072**
Information Entropy	-0.050	-0.021	-0.004	-0.036	-0.056
R-squared Value	0.026	0.039	0.030	0.060	0.022
Video Metadata Variables					
Title Punctuation	0.175***	0.105**	0.093	0.166***	0.181***
Release Time	-0.07*	-0.015	-0.013	-0.029***	-0.067
Video Duration	-0.053	0.26***	0.303***	0.074	-0.046
Video Fan Count	0.045	0.181***	0.247***	0.113	0.051
Increased R-squared	0.032	0.145	0.201	0.106	0.035
Total R-squared	0.058	0.184	0.231	0.166	0.057

Note: Asterisks (*, **, ***) denote statistical significance at the p < 0.10, p < 0.05, and p < 0.01 levels, respectively.

5 DISCUSSION

As online videos continue to be popular, more attention needs paid to the subject content of the videos, rather than just to text information. At the same time, more research are needed to understand the relationship between online video features and video communication effects in order to achieve better communication effects. With the rapid rise of artificial intelligence, using deep learning algorithms to intelligently and automatically extract the main content features in videos and study their relationship with video communication effects has gradually become one of the development directions in the field of computational communication.

We extracted four video feature indicators based on visual frameworks, four video metadata indicators, and four indicators reflecting video communication effects, and investigated the interrelationships among them. Through regression analysis, this study finds several valuable findings.

Firstly, the face appearance rate demonstrated a highly significant positive impact on communication recognition, participation, and social degrees, precisely distinguishing between "viewing" and "engagement." While viewers may click to play a video due to its title or thumbnail, their decision to engage in deep interactions such as giving coins, commenting, or sharing strongly depends on emotional connection and identification with the characters in the video. The individuals in the video serve as "para-familiar" figures; the higher their appearance frequency, the more likely viewers are to establish a one-sided intimate relationship with them. Consequently, viewers become more willing to incur costs (e.g., giving coins) to show support, invest time (e.g., commenting) to interact, or even share the content within their social networks (e.g., forwarding). This demonstrates that the human face acts as a carrier of emotion and trust.

Secondly, in food-related videos, the appearance rate of food has shown a significant negative impact on all communication effect metrics. This counterintuitive finding reveals a shift in the deeper motivations of the audience for such content: their core demand may not be to learn about the food itself, but rather to seek emotional comfort, a sense of companionship, and a social experience—essentially, they are interested in the story between the people and the food.

Furthermore, image brightness demonstrates a significant positive correlation with communication effects. This underscores the importance of fundamental visual experience—bright and well-lit visuals convey a sense of pleasure and freshness, forming a basic threshold for attracting user attention. In contrast, image entropy shows no significant relationship with communication outcomes, suggesting that viewers prefer clear, aesthetically pleasing, and well-composed visuals rather than chaotic or overly complex imagery.

Among the control variables, the use of question marks or exclamation marks in video titles significantly enhances communication effects, highlighting the effectiveness of “clickbait” strategies in stimulating curiosity and emotional responses. Both Video Fan Count and Video Duration emerged as strong predictors of deep engagement. Notably, longer video duration showed a slightly stronger correlation with communication approval and participation than fan count, suggesting that users who commit to watching longer content are also more likely to interact deeply. The significant impact of fan count also reflects the strong Matthew Effect and high fan loyalty within the platform's ecosystem. Conversely, release time shows a negative correlation with communication effects, suggesting that newly published videos are more likely to achieve better dissemination outcomes. This aligns with the internet culture of “fleeting trends”, where attention is consistent drawn to newly emerging content.

This study provides a new perspective on computational methods and communication content in the field of computational communication. The computational method based on big-data, automation, and intelligent recognition, and can applied to other types of video content analysis research directly, such as news, sports, and cute pets. At the same time, since this study focuses on the correlation between the main content of the video and the video dissemination effect, the extracted video feature indicators are relatively small, and more factors that may affect the video communication effect should take into account. In the next stage, we recommend adding more video feature variables to explore the correlation between them and the video communication effect more comprehensively.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

The authors are grateful for the finance support from the Ministry of Education Industry-Academia Collaborative Education Program (No.231104575131914), General Program of National Social Science Fund of China (No. 22BXW040) and Shanghai Jiao Tong University Participation in Research Program (No. T200PRP42015).

REFERENCES

[1] Xiang Y, Chae S W. Influence of perceived interactivity on continuous use intentions on the danmaku video sharing platform: Belongingness perspective. International Journal of Human–Computer Interaction, 2022, 38(6): 573–593. DOI: 10.1080/10447318.2021.1952803.

[2] Song S, Zhao Y C, Yao X, et al. Short video apps as a health information source: an investigation of affordances, user experience and users’ intention to continue the use of TikTok. Internet Research, 2021, 31(6): 2120–2142. DOI: 10.1108/intr-10-2020-0593.

[3] Matei S A, Kee K F. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. Computational Communication Research, 2019, 9(4): e1304. https://tinyurl.com/4ekscyv4.

[4] Trilling D, Araujo T, Kroon A, et al. Computational Communication Science in a Digital Society. Communication Research into the Digital Society Fundam, 2024: 247. DOI: 10.5117/9789048560592.

[5] Joo J, Bucy E P, Seidel C. Computational communication science| automated coding of televised leader displays: detecting nonverbal political behavior with computer vision and deep learning. International Journal of Communication, 2019, 13: 23. https://tinyurl.com/37rcjsav.

[6] Wu P Y, Mebane Jr W R. Marmot: A deep learning framework for constructing multimodal representations for vision-and-language tasks. Computational Communication Research, 2022, 4(1). DOI: 10.5117/CCR2022.1.008.WU.

[7] Jürgens P, Meltzer C E, Scharkow M. Age and gender representation on German TV: A longitudinal computational analysis. Computational Communication Research, 2022, 4(1). DOI: 10.5117/CCR2022.1.005.JURG.

[8] Joo J, Steinert-Threlkeld Z C. Image as data: Automated content analysis for visual presentations of political actors and events. Computational Communication Research, 2022, 4(1). DOI: 10.5117/CCR2022.1.001.JOO.

[9] Zhong Z, Zhou J, Su S. Computational Communication Studies in Chinese and English: Issues, Methods and Theories. Global Journal of Media Studies, 2023, 10(1): 56–76. https://tinyurl.com/yv2jyxpr.

[10] Sun, S., & Liu, Z. Issue Visibility and Gatekeeping Actors in Weibo Trending Searches: A Computational Communication Analysis Based on Public Opinion Events. Journalism & Communication, 2024(5): 30–36+43. https://tinyurl.com/ycyfx7u3.

[11] Sun J. A study on the influencing factors of the dissemination effect of popular science journals on WeChat: a computational analysis based on 4445 tweets. Chinese Journal of Scientific and Technical Periodicals, 2023, 34(11): 1511–1520. DOI: 10.11946/cjstp.202308080600.

[12] Minghua X, Liqun Y, Ziyao W. Hidden Concurrency and Selective Incitement: The Generation Logic and Operational Mechanism of Western Computational Propaganda against China. Contemporary Communication, 2024(3): 55–60. https://tinyurl.com/5arwhdtk.

[13] Qing Y, Yu L. The social support seeking among Depressed patients on social media——Based on the online community of Depression in Weibo. Journalism and Mass Communication, 2022, (6): 45–56+64. DOI: 10.15897/j.cnki.cn51-1046/g2.20220620.004.

[14] Ahmadi I, Waltenrath A, Janze, C. Congruency and users' sharing on social media platforms: a novel approach for analyzing content. Journal of Advertising, 2023, 52(3): 369–386. DOI: 10.1080/00913367.2022.205568.

[15] Li H, Wu Y, Yuan F, et al. Innovative Paths of Computational Communication: A Systematic Analysis of Theoretical Framework, Research Methods and Application Practices——A Review of Computational Communication Research in 2024. Journal of Education and Media Studies, 2025(1): 32–39. DOI: 10.19400/j.cnki.cn10-1407/g2.2025.01.010.

[16] Van Atteveldt W, Peng TQ. When communication meets computation: Opportunities, challenges, and pitfalls in computational communication science. Communication Methods and Measures, 2018, 12(2-3): 81–92.

[17] Yu G, Xu W. Multimodal Fusion: Efficiency Improvement and Research Model of Media Communication. Media Observer, 2021(12): 14–20. DOI: 10.19480/j.cnki.cmgc.2021.12.004.

[18] Huang Y, Chen C. Automation of visual communication and the aesthetic construction of national image: A computational aesthetics study based on Twitter social robots. Modern Communication (Journal of Communication University of China), 2023, 45(8): 96–104. DOI: 10.19997/j.cnki.xdcb.2023.08.016.

[19] Guan L, Zhou B. Application of computer vision technology in news communication research. Contemporary Communication, 2022(3): 20–26. https://tinyurl.com/yc3vh8zw.

[20] Wu Y, Fan J, Zhang L. The Personalization Effect of Mainstream Media in the Context of ShortVideo--Content Analysis of the "Anchor Talking Broadcast" Programme. Journal of Xi'an Jiaotong University (Social Sciences), 2021, 41(2): 131–139. DOI: 10.15896/j.xjtuskxb.202102015.

[21] Xue P. Interactive Mechanism of Online Mukbang Community on Bilibili Video Website. Master's thesis, 2020.

[22] Rodriguez L, Dimitrova D V. The levels of visual framing. Journal of Visual Literacy, 2011, 30(1): 48–65. DOI: 10.1080/23796529.2011.11674684.

[23] Wang J, Chan K C, Loy C C. Exploring clip for assessing the look and feel of images. In Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(2): 2555–2563. DOI: 10.1609/aaai.v37i2.25353.

[24] Yu W. Did the "Shannon-Weaver" Model Ever Exist?——Critical Examination of Key Historical Facts in Domestic and International Communication Theory from an Interdisciplinary Versioning Approach. Journal of Journalism & Communication Studies, 2024, 31(10): 19–37+126. https://tinyurl.com/aadxnhh9.